基於深度學習之平面影像深度化與無人車

一、研究計畫內容（以10頁為限）：

1. 摘要

深度學習(Deep Learning)的概念在1992年左右正式被提出，但受限於當時得晶片技術，只有超級電腦可以負擔如此龐大的運算，但隨著GPU浮點運算能力的大幅進步，小型實驗室甚至個人皆可以輕鬆取得這些運算能力，深度學習才算是真正進入世人的眼睛，也勢必大幅改變我們生活模式。DL(Deeping Learning)為機器學習底下的一個分支，我們透過把資料送進多個處理層(layer)中進行縣性或非線性轉換，抽取足以代表資料特性的特徵。早期，我們需要使用人力進行資料分析研究，以產生良好的特徵，這便是所謂的特徵工程(Feature Engineering)。深度學習的引入，可以達到自動抽取特徵的目的，被視作一種特徵學習(Feature Learning)，大幅縮減特徵工程所花的人力與時間。隨著深度學習日趨成熟，發展出各式各樣的演算法，其中也包含經由卷積神經網路(CNN)來測量距離。如果我們能運用深度學習在在單鏡頭取代以往無人車雙鏡頭的測距功能，那勢必可以節省一部份成本，並且讓自駕車這塊領域更為普及化。

1. 研究動機與研究問題

現今的無人車在製造時仍是需要相當多的成本，於是我們在討論主

題時就考慮了多種可能，也查詢了相當多的資訊，最後看到無人車一般

製作時需雙鏡頭已達測距的功能，我們就想那如果將其改成單鏡頭呢?

這便構成了我們的研究動機。

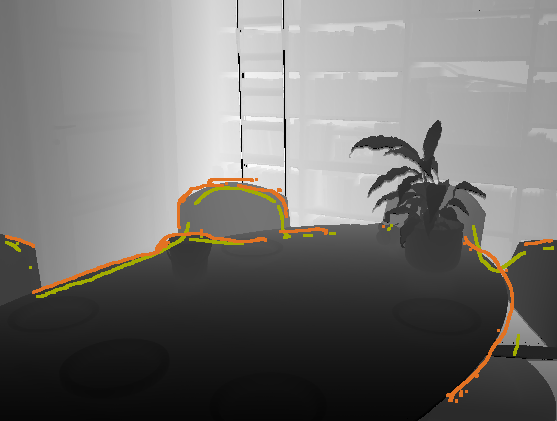
自駕車在避障時需要高解析度的前方3D影像，偵測物體的大小規劃合適路徑，業界大多使用雙鏡頭使用相位差的方式進行影像深度判斷，但市面上雙鏡頭的成本相當高昂，約是單鏡頭的3倍到4倍左右，造成一個相當大的門檻，因此我們希望使用開發一種低成本的深度影像感測技術，使開發小型無人車可以是負擔的起的，並且我們將實際做出一台以此為感測元件的無人車，將利用學校環境作為測使場地，在固定路徑上蒐集以雙目鏡頭為主的無人車以及以單目鏡頭深度影像化的無人車在相同路徑上行駛需要使用的時間，來測試結果。我們使用單鏡頭運用卷積神經網路(CNN)以判斷障礙物的距離，隨著近年來神經網路的快速發展，也有相當多人在卷積神經網路做很多的研究，利用卷積層層堆疊的特性，以求得精確的數值。如果我們達成這一步，勢必可以減去發展自駕車部分成本，帶動其技術更為成熟。無人車的使用將會搭載ROS (Robotic Operating System)作為主控系統，本系統將會參照美國麻省理工學院在2016年所做的MIT-Racecar的研究，他使用的無人車是搭載ZED Stereo Camera雙目鏡頭作為探測距離的工具，我們將會以單目鏡猴取代，並比較成果。



圖一、MIT-Racecar 成品

1. 文獻回顧與探討
2. Evaluation of CNN-based Single-Image Depth Estimation Methods

隨著近年來深度學習領域對計算視覺領域有著極大的影響，SIDE的問題也得到相當多的解答。然而為了增進傳統的方法，卷積神經網路(CNN)以變最常用且能從單張圖片中得知相當多的訊息，可惜的是，在以往對於室內深度的數據集較為匱乏，顯示出對於邊界周圍有著明顯缺陷，且同時阻礙了精確的評估。於是我們做了一個新的評估數據集，包含高精度的RGB圖片，可運用在評估各種室內數據及性能測試。



(a)ground truth (b)predicted

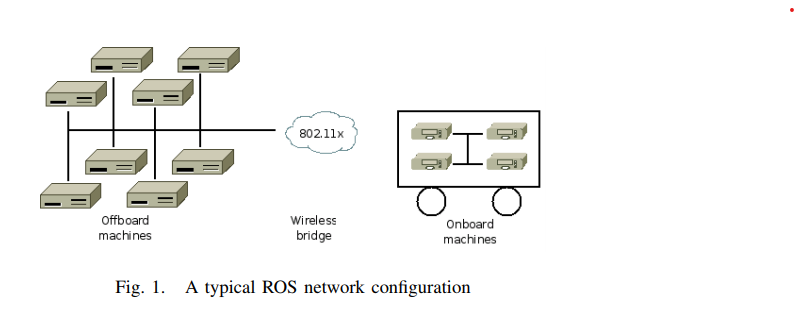


(c)depth plane (d)different

圖1、運用DBE(a+b)和DDE(c+d)所得出的結果，當中橘線是利用Li的方法，而綠線是利用Laina方法，而藉由c、d可得知利用Li方法對比真實深度，橘色區塊為較近，而綠色區塊為較遠。

1. ROS: an open source Robot Operating System

這篇論文給予了一個ROS(Robot Operating System)的概觀，ROS並不是傳統意義上程序管理的作業系統，由各大學所共同開發，目的是為了在開發新機器人技術實不必從解決繁瑣的硬體軟體問題，可以直接運用其他大學所完成的部分，以其為基礎將時間全然用於開發新技術上，ROS提供了一個在電腦群集底下的結構化通信層，將不同的功能個別分散於各個節點，並且運用P2P(Peer-to-Peer)網路進行資料傳輸，下達指令。



1. 研究方法及步驟
2. 甘特圖

11月 方向確定、文獻蒐集、撰寫研究動機目的

12月 購入器材、學習相關知識、熟悉程式搭載環境

1月 開始撰寫摘要(研究動機、文獻回顧、研究方法、預期結果)及其餘內容

2月 完稿

3月 完成申請

1. 研究方法及步驟

我們將分為兩部分同時進行

第一部分：單鏡頭深度深度影像

(1)卷積層(convolution Layer)

藉由輸入不同的特徵，一層一層從低階特徵疊加上去。每

層卷積層都由若干卷積單元所組成。

(2)池化層(Pooling Layer)

將輸入的圖像化分為若干區塊，將各個區塊進行池化，過

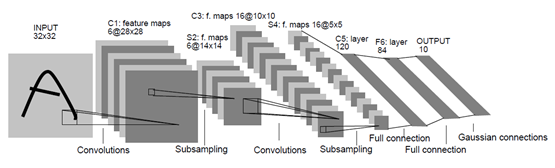
去常用平均池化(Mean Pooling)但是後來發現最大池化

(Max Pooling)的效果較好，所以現階段平均池化已較少

用。

(3)全連接層(Fully Connected Layer)

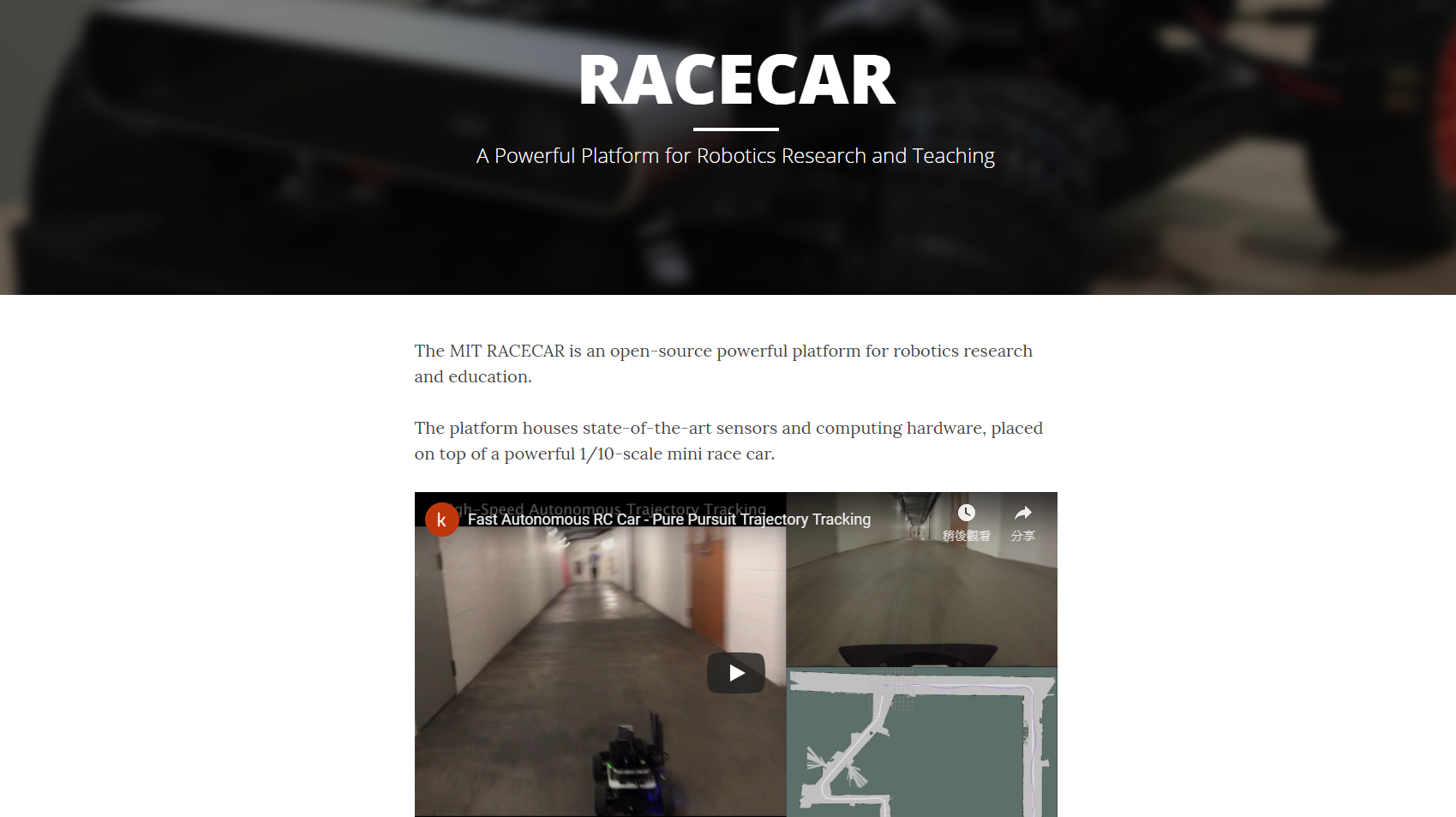
相當於一般的神經網路(MLP)



圖二、捲積神經網路示意圖(此為lenet-5，取自於Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick ´Haffner. Gradient-based learning applied to document recognition.)

第二部分：小型化無人車

我們參考MIT Racecar進行無人車的設計，MIT-RACECAR是麻省理工學院一個強大的開放源碼的機器人研究和教學平臺。該平臺內置最先進的感測器和計算硬體，放置在功能強大的1/10規模迷你賽車上。RACECAR設計和開發是林肯實驗室的BeaverWorks計畫，航空航天部門和麻省理工學院資訊和決策系統實驗室之間的共同完成。



(1)購買的硬體配置有

a.車載嵌入式開發板：NVIDIA Jetson TX2

b.遙控車底盤：TRAXXAS 1/10 Scale 4-Tec 2.0 VXL AWD

Chassis

c.2D光學雷達：RPLidar A2

d.雙目鏡頭：ZED Stereo Camera

e.IMU慣性測量單元：RAZOR 9DOF IMU

f.電子調速器：VESC

g.WiFi路由器

(2)安裝ROS機器人作業系統於Jetson TX2上：

ROS是一個Open source的機器人開發平台，在這個架構下，讓機器人各大領域Navigation、Manipulation、Perception、Cognition等，建立共通的使用平台，全世界的研究員都可以在這個平台上分享各種機器人相關研究的資源和演算法

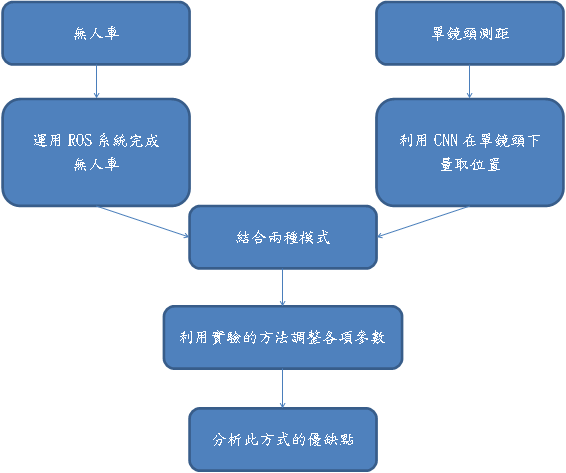
　　要做一台機器人，有ROS的話就不需要要從頭寫Controller、Driver甚至是3D物理環境模擬器。ROS提供我們一個穩定的模擬測試平台，以及共用各知名大學的研究成果，讓我們可以專注在改善機器人不足的地方，能夠大力發展各大學的專長技術。

　　目前ROS已支援各式各樣的程式語言，已實做且驗證穩定的的包括C++和Python；實驗性的程式庫，也就是已經提供該種程式語言相對應的Client Library。

(3)安裝MIT-Racecar既存開發完成的部分，進行硬體調校，最終使其能在雙目鏡頭下施行無人駕駛。

第三部分：無人車與單鏡頭深度影像的結合

以上兩部分完成後，會開始將深度資料使用單鏡頭深度演算提供，並進行實測，在固定路徑上與使用雙鏡頭比較，檢驗成效。



1. 預期結果

我們預期能順利運用單鏡頭達成景象深度的機械學習，並且能實際運用到無人車上，且能在室內教室完成試跑加上順利避障的功能。

現今深度學習的發展越來越成熟，也誕生了各種不同的演算法，如CNN、RNN、Alexnet、Lenet…等等。我們希望能找出最為合適的架構，並且能順利地運用在我們的預期，精準的分類和測量，從中找出其優缺點，使得這項技術能更加成熟和被廣泛地運用，這是我們最終的目標。

1. 參考文獻
2. Tobias Koch1 Lukas Liebel1 Friedrich Fraundorfer2,3 Marco Körner

Evaluation of CNN-based Single-Image Depth Estimation Methods

1. ROS: an open-source Robot Operating System Morgan Quigley∗ , Brian Gerkey† , Ken Conley† , Josh Faust† , Tully Foote† , Jeremy Leibs‡ , Eric Berger† , Rob Wheeler† , Andrew Ng∗ ∗Computer Science Department, Stanford University, Stanford, CA †Willow Garage, Menlo Park, CA ‡Computer Science Department, University of Southern California
2. 需要指導教授指導內容

1.影像辨識

2.機器學習